

Volume 9

1950

PREMODULATION SPEECH CLIPPING AND FILTERING
A CONSIDERATION OF THEIR EFFECTS ON THE
INTELLIGIBILITY OF SPEECH

by

Stephen E. Stuntz

BU MEDICAL CENTER
FILE COPY NO. 155
PP. 99-96

Medical Research Laboratory Report No. 155
BuMed Project NM 003 041.21.04

APPROVED: T. L. Willmon, Captain (MC) USN; OinC 14 August 1950

PREMODULATION SPEECH CLIPPING AND FILTERING:
A CONSIDERATION OF THEIR EFFECTS ON THE
INTELLIGIBILITY OF SPEECH

by

Stephen E. Stuntz

Reference may be made to this report as follows:
S. E. Stuntz, Premodulation Speech Clipping and
Filtering: A Consideration of their Effects on the
Intelligibility of Speech, MRL Report No. 155, 9,
90-96, (1950).

BuMed Project NM 003 041.21.04

ABSTRACT

Speech clipping and filtering circuits in voice communication systems so distort the transmitted speech as to raise a question of their effect on intelligibility. Results of several experiments indicate that speech clipping alone improves intelligibility when signals are partially masked by electrical noise entering the system between transmitter and receiver. Signals which are not so masked are similarly affected to almost the same degree. Intelligibility improves as clipping increases up to at least 24 db. Under extreme noise masking, very heavy clipping (100 db or more) improves intelligibility, but reduces intelligibility of signals not so masked. Other experiments show that intelligibility is not likely to suffer when voice frequencies below 580 cycles are sharply attenuated, with the possible exception that when masking white noise is very intense, restoring low frequencies causes a slight increase of intelligibility. Attenuating frequencies above 3900 cycles has no significant effect on intelligibility; however, reducing the low-pass cut-off point to 2500 cycles does lower intelligibility considerably. It is suggested that band-pass filtering limits be 600 to 4000 cycles. A list of references is appended.

PREMODULATION SPEECH CLIPPING AND FILTERING: A CONSIDERATION OF THEIR EFFECTS ON THE INTELLIGIBILITY OF SPEECH

Over the past ten years, the radio journals have published extensively on amplitude- and frequency-limiting speech amplifiers. The fundamental purpose of such equipment has been the improvement of transmission efficiency in radiophone transmitters, by permitting only the frequencies essential to speech to modulate the transmitter, and by maintaining the highest average proportion of modulation possible without exceeding the transmitter's linear capability. The principal devices recommended have used volume compressors, limiters, clippers, or automatic gain-control networks, and usually incorporate frequency filters which eliminate all but the band occupied by typical speech.

The physics and economics of these amplitude- and frequency-restricting amplifiers have apparently been generally accepted. However, there seems to have been a good deal of unwillingness to employ them, regardless of their technical advantages, on the ground that they so distort the natural voice as to jeopardize intelligibility. This argument does not square with the results of some recent experiments designed to study exactly this problem: how do clipping and filtering affect speech intelligibility? Indicative of this work was an article published over two years ago in QST¹. Several others have appeared in the engineering and scientific literature which are worth the amateur phone man's consideration. This paper will undertake a review of these articles, and attempt to demonstrate the usefulness of their results.

First, let us refresh our knowledge of the dynamics of speech. When the human voice is impressed upon a microphone, voltages are set up whose instantaneous peaks normally exceed the root-mean-square value by 12 to 15 db². It is this "peak factor" which requires us to design our amplifiers with a much greater range of linear amplification than we expect to use most of the time, when we wish to transmit voice signals with minimum distortion. Also, in English as we speak it, the average vowel produces a peak voltage which runs about 12 db higher than that produced by the average consonant^{3,4}. This 12 db figure is the average vowel-consonant ratio for all combinations of sounds in our language; the instantaneous value may vary from a fraction of the average to several times its magnitude. Now, peculiarly enough, the intelligibility of speech depends much more

heavily upon the sounding of consonants (b, p, z, s, t, d, f, v, th, k, l, m, n, etc.) than upon vowels (a, e, i, o, u, y, etc.), despite the fact that the ordinary vowel sound has around 16 times the power of the usual consonant.

Effects of Clipping* on Intelligibility

From the foregoing, we can immediately see what happens when the peaks are clipped from the speech wave. At one and the same time we reduce (1) the peak factor, and (2) the vowel-to-consonant ratio. Effectively, we have cut down the range of variation in speech-energy amplitude, and in so doing have given proportionally greater emphasis to consonants, upon which intelligibility largely depends, as we have seen. This would lead us to expect that we might improve intelligibility by the use of clipping. There is, on the other hand, the possibility that the distortion of amplitudes resulting from peak-clipping might actually reduce intelligibility. This is the gist of the question for which answers have been sought in the psychological laboratory, using some techniques of measurement which have become standard in studying voice communication.

Some years ago the Bell Telephone Laboratories devised tests to measure the effects of telephone circuits on the intelligibility of speech. A talker would read lists of syllables or words made up of all the sounds of the English language, in various combinations. His voice was then transmitted over a telephone circuit to a group of listeners who would write down what they thought the talker had said. By comparing the talker's original list with the listeners' reproduction of it, a percentage could be computed representing the proportion of spoken sounds correctly received by the listeners⁵, as circuit conditions were systematically changed by introducing various degrees of filtering, attenuation, nonlinear amplification and the like. Early in World War II, this method was applied by a group of psychologists at the Harvard University Psycho-Acoustic Laboratory, to a study of the effects of premodulation clipping upon the intelligibility of speech transmitted over a miniature radiophonic circuit, using standard amplitude modulation⁴. The results of this testing show that in the absence of QRN⁶, when extremely weak unclipped signals were only about 30% intelligible,

*In this discussion as in general engineering usage, "clipping" means cutting off both positive and negative speech-wave peaks to an equal extent, then reamplifying the resultant up to the magnitude it would have attained without clipping.

⁴In the international code of telegraph operating signals, the letter group QRN is used in reference to the noises produced by atmospheric static. It is used in this sense throughout this paper as a matter of convenience.

using 24 db of clipping would raise intelligibility to 75%. These percentages represent intelligibility of words on the special lists; the equivalent in connected meaningful sentences may be higher^{3,5}. Such an advantage in favor of clipping also holds when QRN is very heavy, to almost the same degree. Listeners in these experiments were asked to report on change of voice quality as clipping increased. Here is their average opinion: at 0 db clipping, natural voice; 6 db, clipping effects barely noticeable (comparable to standard broadcast quality); 12 db, talker appeared to be enunciating with unusual care; 18 db, voice took on a sharp "sandy" character, quality rated not as good as before; 24 db, voice was coarse and "grainy", rated as poor. Note, however, that despite the very evident changes in voice-quality, intelligibility actually improved, particularly when conditions were less than optimal. This effect has been noted before¹, although perhaps not so explicitly documented.

The question arises, "What about the effects of clipping on intelligibility when conditions are nearly perfect?" The most definitive answer available comes from another series of experiments⁶. With 0 db clipping, signals were 100% intelligible; as clipping was increased, intelligibility fell-off slightly until at 20 db clipping it had reached 96%. Clipping was gradually advanced, and at 100 db (almost all speech peaks flattened to rectangles), intelligibility had fallen to 75%. (This, remember, under ideal conditions of quiet for both talker and listeners, with no fading or interference.) Incidentally, these experiments revealed that with signal/noise conditions which completely obscured unclipped speech (intelligibility at 0% to 10%, the same signal when clipped 100 db and over was 30% intelligible). In these experiments, nothing was said about changes in quality; it could be expected, however, that with such severe clipping as 100 db it would be very hard to identify the talker by the distinctive sound of his voice.

Effects of Filtering on Intelligibility

Up to this point, we have discussed speech in terms of its gross amplitudes only, without considering the individual frequencies present in spoken language. For reasons dictated by engineering standards, several recent amplifier designs^{7,8,9} have included both high- and low-pass filtering. Since this practice is becoming more widespread, let us examine its effect on intelligibility.

This matter has also been investigated in the psychological laboratory, under conditions comparable to those found in the amateur phone bands^{10,11}. Using the same testing procedures as in the study of clipping effects, a talker's voice was transmitted over a wire circuit to a group of listeners. The speech was subjected to various degrees of filtering and attenuation, and was then combined with an unfiltered, constant-intensity thermal noise, simulating QRN, and led to the listeners' headphones. At no time was peak-clipping permitted to occur; thus the effects of filtering alone could be evaluated. One series of experiments, studying changes of intelligibility at various signal levels and signal noise ratios when either high or low frequencies were separately filtered out¹⁰, showed that when everything below 350 cycles was cut off, intelligibility of moderately to very strong signals was slightly improved by comparison with unfiltered signals. However, at the lower signal levels, where QRN presumably was more disturbing, intelligibility suffered loss as a result of such filtering. Extremely weak signals in noise were 5% intelligible when the 350-cycle high-pass filter was in the circuit, but jumped to 25% when the filter was switched out, although signal strength and noise level remained unchanged. It was further found that when signals were strong and in the clear, everything up to 580 cycles could be cut off with little damage to intelligibility. As to cutting off the highs, when everything above 3950 cycles was eliminated there followed very little reduction of intelligibility regardless of signal strength. However, when the cutoff point was moved down to 2500 cycles, results were quite different. When signals were strong and clear, intelligibility was down to 78% with the filter in, as compared to 90% with no filter. As signals grew weaker, the proportional loss of intelligibility due to filtering diminished somewhat, although even at the lowest signal level used in the tests, the 2500-cycle low-pass filter hampered intelligibility appreciably.

We may now ask, "What happens to intelligibility when we filter off both highs and lows at the same time?" The effects of band-pass filtering on speech in a noise background have been separately investigated¹¹. As before, unfiltered constant-intensity noise was superimposed upon the filtered speech signal, which was also varied in strength to secure various signal/noise ratios. As might be expected from the discussion of high- and low-pass filtering, greatest intelligibility at all signal strengths resulted when the widest pass-band was used (130-9200 cycles, intelligibility about 90%).

The effects of filtering upon intelligibility were most noticeable, as before, when signals were strong and relative noise-level was low. Interestingly enough, at all signal-levels, the pass-bands 340-3900 cycles and 550-3900 cycles produced almost identical effects on intelligibility; actually, neither one seriously impaired intelligibility when compared with the widest pass-band. However, shifting the cutoff points toward each other clearly resulted in poorer intelligibility, as the following table shows. Signal-strength and signal noise ratio are the same for all filter combinations.

Pass-band limits	Intelligibility
130-9200 cycles	90%
340-3900	80
550-3900	80
550-2500	70
870-3900	65
870-2500	55

There have been several other investigations of filtering and its effect upon speech intelligibility^{3,12}. However, they deal only with signals transmitted under ideally quiet, unfading conditions for both talker and listeners; their results, therefore, are probably not as pertinent to amateur phone operation as the work already discussed, hence no detailed mention of them has been made.

Summary

On the basis of the foregoing evidence, we may summarize as follows:

1. Speech-clipping definitely improves intelligibility.
2. As signals get weaker, and as signal/noise ratio gets worse, the greater the clipping, the greater the improvement of intelligibility, up to at least 24 db of clipping.
3. Extremely heavy clipping (100 db or more) is beneficial under very severe signal/noise conditions, although it will probably not make poor signals completely understandable.
4. Although the quality of speech changes noticeably over the clipping range from 0 db to 24 db (and probably above), even under the best signal conditions intelligibility

is not impaired by clipping.

5. In general, high-pass filtering up to 350 cycles will not harm intelligibility, and may actually make a slight improvement when signals are strong and clear.

6. Under optimum signal conditions, frequencies below 580 cycles may be eliminated with little loss of intelligibility.

7. Cutting off frequencies above 3900 cycles by use of a low-pass filter will have hardly any effect on intelligibility.

8. Cutting off frequencies above 2500 cycles will seriously impair intelligibility.

Conclusions

We may conclude, therefore, that the engineering advantages obtained from speech clipping prior to modulation are accompanied by definite improvement of intelligibility at the receiving end of a radio circuit; especially under adverse operating conditions. Further, the change in voice-quality noted as a by-product of clipping does not really impair intelligibility of the signal; speech can be distorted very severely by nonlinear transmission and still be perfectly understandable. Filtering to avoid or remove the undesirable side-effects of clipping will not impair the intelligibility of speech until the upper cutoff frequency gets down around 2500 cycles. In fact, filtering off the low frequencies (below 350) may actually improve intelligibility under good signal conditions. The limit for cutting off low frequencies is apparently much less critical than for high frequencies: any cutoff point up to almost 600 cycles may serve for the lows with little damage to intelligibility, while for the highs cutoff should be well above 2500. It appears now that the more or less arbitrary low-pass cutoff of 3000 cycles now rather widely employed^{7,8,9} may be a little too low for optimum communication. This last observation assumes, of course, that the frequencies above nominal cutoff are abruptly and completely attenuated. It may very well be that intelligibility would not suffer so seriously were the frequencies above, say, 2000 cycles subjected to the relatively gentle treatment of the typical R-C network, i.e., 3 to 6 db attenuation per octave.

REFERENCES

1. Smith, W. V. Premodulation speech clipping and filtering. *QST Magazine*, 30, 2, Feb., 1946, 46-50.
2. Dunn, H. K., and White, S. D. Statistical measurements on conversational speech. *Journal of the Acoustical Society of America*, 11, 3, Jan., 1940, 278-288.
3. Fletcher, Harvey. *Speech and Hearing*. Van Nostrand, Inc., N. Y., 1929.
4. Kryter, K. D., Licklider, J. C. R., and Stevens, S. S. Premodulation clipping in AM voice communication. *Journal of the Acoustical Society of America*, 19, 1, Jan., 1947, 125-131.
5. Egan, J. P. Articulation testing methods. *Psycho-Acoustic Laboratory, Harvard University, OSRD Report No. 3802*, Nov. 1, 1944, PB 22848.
6. Licklider, J. C. R., Bindra, Dalbir, and Pollack, Irwin. The intelligibility of rectangular speech waves. *American Journal of Psychology*, 61, 1, Jan., 1948, 1-20.
7. Rand, P. S. The "little slugger". *QST Magazine*, 33, 2, Feb., 1949, 11-17.
8. Anon. Logarithmic compressor. *G. E. Ham News*, 5, 3, May-June, 1950.
9. Swafford, T. W. More effective speech amplification. *QST Magazine*, 34, 7, July, 1950, 50-52.
10. Pollack, Irwin. Effects of high pass and low pass filtering on the intelligibility of speech in noise. *Journal of the Acoustical Society of America*, 20, 3, May, 1948, 259-266.
11. Egan, J. P., and Wiener, F. M. On the intelligibility of bands of speech in noise. *Journal of the Acoustical Society of America*, 18, 2, Oct., 1946, 435-441.
12. French, N. R., and Steinberg, J. C. Factors governing the intelligibility of speech sounds. *Journal of the Acoustical Society of America*, 19, 1, Jan., 1947, 90-119.